

## 2

### Modelo em Espaço de Estado

#### 2.1

##### Idéia básica

O Modelo em Espaço de Estado (MEE) proporciona uma metodologia unificada para o tratamento dos problemas em séries temporais. Nesta abordagem, assumimos que o desenvolvimento do sistema no tempo em estudo é determinado por uma série de vetores  $\alpha_1, \dots, \alpha_n$  não observados associados as séries das observações  $y_1, \dots, y_n$ . A relação entre  $\alpha$ 's e  $y$ 's é especificada pelo MEE.

Na literatura de séries temporais macroeconômicas, uma particular estrutura em espaço de estado é utilizada, conhecida como Modelos Estruturais. Nesta abordagem o vetor de estado é decomposto em componentes tais como tendência, inclinação, sazonalidade e ciclo. Desta forma, é possível extrair cada um dos componentes no processo, os quais possuem interpretação econômica direta.

#### 2.2

##### Modelos Gaussianos e Lineares

O modelo em espaço de estado Gaussiano linear geral pode ser escrito como:

*-Equação das observações:*

$$y_t = Z_t \alpha_t + \epsilon_t \quad \epsilon_t \sim N(0, H_t),$$

*-Equação do estado:*

$$\alpha_{t+1} = T_t \alpha_t + R_t \eta_t \quad \eta_t \sim N(0, Q_t),$$

$$t = 1, \dots, n \quad (2.1)$$

onde  $y_t$  é um vetor  $p \times 1$  das observações e  $\alpha_t$  é um vetor não observável  $m \times 1$  chamado de vetor de estado;  $R_t$  é a matriz de seleção de  $\eta_t$  composta por  $r$  colunas da matriz identidade de tamanho  $m$ ;  $Z_t$  é o vetor de seleção  $1 \times m$  do estado;  $T_t$  é a matriz de transição do estado de dimensão  $m \times m$ ;  $H_t$  e  $Q_t$  são as matrizes de covariância de dimensão  $p \times p$  e  $r \times r$  dos distúrbios da observação e do estado, respectivamente. Alguns elementos das matrizes  $Z_t, H_t, T_t, R_t$  e  $Q_t$  dependerão do vetor de parâmetros não conhecido  $\Psi$ , conhecido como vetor de hiperparâmetros.

A primeira equação de (2.1) é chamada de equação da observação. Tem estrutura de um modelo de regressão linear onde o vetor coeficiente  $\alpha_t$  varia no tempo. A segunda equação é chamada equação de estado, que representa um modelo autoregressivo vetorial de primeira ordem. A natureza Markoviana fica caracterizada na equação do estado, consistindo numa propriedade elegante dos MEE.

Levando em consideração as propriedades de otimalidade dos estimadores de máxima verossimilhança, a estimação dos hiperparâmetros dá-se maximizando a função de verossimilhança das observações usando algoritmos de otimização. Para o cálculo da função de verossimilhança é necessário a estimação da densidades *a priori* e *a posteriori* do vetor de estado. No caso Gaussiano as densidades serão Gaussianas com média e variância estimadas pelo filtro de Kalman.

Existem outros métodos de estimação de hiperparâmetros, como o algoritmo EM. É um método iterativo para obtenção de estimadores de máxima verossimilhança em situações onde há dados incompletos. É dividido em dois passos: Estimação (E-step) e Maximização (M-step). Mais detalhes do método no contexto dos MEE em Shumway and Stoffer[41].

A principal vantagem do método EM é que as equações do algoritmo têm formas mais simples que as equações não lineares de outros métodos. A principal desvantagem é a convergência lenta nos últimos passos do algoritmo. Em termos práticos, pode-se utilizar os valores provenientes do algoritmo EM em um determinado número de iterações para inicializar o algoritmo de máxima verossimilhança. Tal procedimento tende a tornar a

busca pelos valores dos hiperparâmetros mais eficiente do ponto de vista computacional.

### 2.2.1

#### Filtro de Kalman

O filtro de Kalman[14] é um elemento essencial para a análise estatística dos MEE Gaussianos. Calcula de forma recursiva o estimador da média e covariâncias da densidade a *posteriori* do vetor de estado conhecido o vetor de hiperparâmetros  $\Psi$ . Em Harvey[22] prova-se que, em ambiente Gaussiano, o filtro de Kalman produz o melhor estimador linear no sentido de minimizar o erro quadrático médio.

A partir da premissa de Gaussianidade dos distúrbios e das condições iniciais do modelo, as expressões recursivas podem ser obtidas. Utilizando os resultados clássicos da distribuição normal multivariada, calcula-se de forma recursiva a distribuição condicional do vetor de estado  $\alpha_t$  condicional às observações até o tempo  $t$ . Tal distribuição também será Gaussiana, e portanto, determinada pela média e variância calculadas pelo filtro de Kalman[14, 20, 22].

Sejam  $a_{t/t-1} \hat{=} E[\alpha_t/Y_{t-1}]$ ,  $P_{t/t-1} \hat{=} Var[\alpha_t/Y_{t-1}]$  a média e variância, respectivamente, da distribuição a *priori* do vetor de estado;  $a_{t/t} \hat{=} E[\alpha_t/Y_t]$ ,  $P_{t/t} \hat{=} Var[\alpha_t/Y_t]$  a média e variância, respectivamente, da distribuição a *posteriori* do vetor de estado;  $\hat{y}_{t/t-1} \hat{=} E[y_t/Y_{t-1}]$  a previsão um passo à frente e as inovações como sendo  $v_t$ . O filtro de Kalman é dado por:

Inicialização:

$$a_{1/0} \text{ e } P_{1/0}$$

Inovação:

$$v_t = y_t - Z_t' a_{t/t-1}$$

Ganho de Kalman:

$$K_t = T_t P_t Z_t' F_t^{-1}$$

$$Var(y_t/Y_{t-1}) = F_t = Z_t P_{t/t-1} Z_t' + H_t$$

Previsão do Estado:

$$a_{t/t-1} = T_t a_{t-1}$$

$$P_{t/t-1} = T_t P_{t-1} T_t' + R_t Q_t R_t'$$

Atualização do Estado:

$$a_{t/t} = a_{t/t-1} + P_{t/t-1} Z_t' F_t^{-1} v_t$$

$$P_{t/t} = P_{t/t-1} - P_{t/t-1} Z_t' F_t^{-1} Z_t P_{t/t-1}$$

$$t = 1, \dots, T$$

Deste modo, o filtro de Kalman fornece média e variância condicionais do estado. O que permite a construção da densidade preditiva, a qual viabiliza a avaliação da verossimilhança, e conseqüentemente, obtenção das estimativas dos hiperparâmetros via algoritmos de otimização.

Para modelos com componentes não estacionárias a distribuição inicial do vetor de estado  $\alpha_{1/0}$  não é conhecida, assume-se que para o instante inicial, o vetor de estado segue uma *priori* difusa, ou seja,  $\alpha_{1/0} = 0$ ,  $P_{1/0} = \kappa I$ , onde  $\kappa$  é um número grande o suficiente[21]. Este procedimento pode gerar instabilidade numérica no sistema. Em função disso, existem outras soluções para o problema da inicialização como em Rosenberg[38], de Jong[10, 11] e mais recentemente a proposta apresentada por Durbin e Koopman[14].

### 2.2.2

#### Filtro de Kalman suavizado

O objetivo do filtro de Kalman suavizado[1], (FKS), é obter estimativas utilizando a totalidade das observações. Desta forma, aumenta-se a acurácia dos parâmetros estimados, o que é interessante quando se analisa o comportamento das componentes não observáveis individualmente. Ao longo do tempo o FKS foi aperfeiçoado[10, 29, 30], conseguindo aumentar a eficiência do algoritmo e diminuir o tempo computacional. A seguir apresentamos uma breve descrição do filtro.

Inicialização:

$$r_T = 0 \quad N_T = 0$$

De  $t=T, \dots, 1$ , temos:

$$\hat{\epsilon}_t = H_t u_t, \quad \text{Var}(\epsilon_t/y) = H_t - H_t (F_t^{-1} + K_t' N_t K_t) H_t,$$

$$\hat{\eta}_t = Q_t R_t' r_t, \quad \text{Var}(\eta_t/y) = Q_t - Q_t R_t' N_t R_t Q_t,$$

$$u_t = F_t^{-1}\nu_t - K_t'r_t, \quad D_t = F_t^{-1} + K_t'N_tK_t,$$

$$r_{t-1} = Z_t'u_t + T_t'r_t \quad \text{e} \quad N_{t-1} = Z_t'D_tZ_t + T_t'N_tT_t - Z_t'K_t'N_tR_t - T_t'N_tK_tZ_t.$$

Da recursividade podemos extrair o vetor de estado suavizado

$$\hat{\alpha}_{t+1/T} = T_t\hat{\alpha}_{t/T} + R_tQ_tR_t'r_t, \quad t = 1, \dots, T.$$

Uma das aplicações desta rotina é o cálculo dos resíduos auxiliares, que são utilizados na investigação de *outliers* e estruturas do sistema. Em nosso contexto o filtro de Kalman suavizado servirá como base do procedimento de suavização por simulação[12]. Esta abordagem é usada no conjunto dos modelos modelos não Gaussianos apresentados à frente.

### 2.2.3

#### Suavizador por simulação

Construído por de Jong & Shephard[12], o método de simulação é eficiente para amostragem das densidades associadas aos modelos não Gaussianos/não lineares para séries temporais. Este método tem aplicação direta nos métodos de MCMC.

Ao invés de simular o vetor de estado, este procedimento produz simulações dos distúrbios aleatórios. Isto é possível porque a quantidades aleatórias são combinações lineares dos distúrbios, e portanto, podem ser extraídas a partir dos mesmos.

De posse das quantidades  $\nu_t, F_t, K_t$ , provenientes do filtro de Kalman e sabendo que:

$$E[\epsilon_t, \eta_t] = G_t, \quad J_t = H_tK_t'N_tG_t,$$

$$L_t = (T_t - K_tZ_t), \quad C_t = Var[\epsilon_t/y] + H_t - H_tD_tH_t - G_t'N_tG_t + J_t + J_t'.$$

Apresentamos aqui uma forma particular do método, onde o distúrbio é sorteado, a cada iteração, da densidade condicional do distúrbio dado todas as observações, ou seja, sorteia-se  $\epsilon^{(i)}$  de  $p_G(\epsilon/y)$ <sup>1</sup>:

$$\epsilon_t^{(i)} = H_t\tilde{\epsilon}_t + G_t\tilde{r}_t + u_t^{(i)}, \quad u_t^{(i)} \sim N(0, \tilde{C}_t),$$

onde as quantidades  $\tilde{\epsilon}_t$  e  $\tilde{C}_t$  são obtidas de

$$\tilde{\epsilon}_t = F_t^{-1}\nu_t - K_t'\tilde{r}_t,$$

<sup>1</sup> $p_G(\epsilon/y)$  é conhecida como densidade de importância, mais detalhes em §2.3.3.

$$\tilde{D}_t = F_t^{-1} + K_t' \tilde{N}_t K_t,$$

$$J_t = J_t = H_t K_t' \tilde{N}_t G_t,$$

$$M_t = H_t (\tilde{D}_t Z_t - K_t' \tilde{N}_t T_{t+1}) + G_t' \tilde{N}_t L_t,$$

$$\tilde{C}_t = H_t - H_t \tilde{D}_t H_t - G_t' \tilde{N}_t G_t + J_t + J_t'$$

e as recursões inversas:

$$\tilde{r}_{t-1} = Z_t' F_t^{-1} \nu_t - M_t \tilde{C}_t^{-1} u_t^{(i)} + L_t' \tilde{r}_t,$$

$$\tilde{N}_{t-1} = Z_t' F_t^{-1} Z_t + M_t \tilde{C}_t^{-1} M_t + L_t \tilde{N}_t L_t,$$

para  $t = T, \dots, 1$  com  $\tilde{r}_T = 0$  e  $\tilde{N}_T = 0$ .

## 2.3

### Modelos Não-Gaussianos e Não-Lineares

Até aqui, foi apresentado como tratar de MEE Gaussianos e lineares. Métodos baseados nestes modelos geralmente conseguem resolver os problemas encontrados em séries temporais típicas encontradas em Economia. Entretanto, existem situações onde o modelo Gaussiano linear falha ao tentar reproduzir o comportamento dos dados. Por exemplo, uma série de dados referentes à retornos financeiros, onde existem evidências de aglomerados de volatilidade e caudas pesadas na distribuição de probabilidade. Seria muito mais apropriado modelar a série com uma distribuição de caudas pesadas, como *t-Student*, do que com uma distribuição Gaussiana. Para esses casos apresentamos uma representação geral que seja capaz de capturar o comportamento dos dados.

### 2.3.1

#### Modelo geral não Gaussiano

O modelo geral multivariado não Gaussiano que será apresentado tem estrutura similar a estrutura espaço de estado em (2.1) onde os vetores das observações  $y_t$  são determinados pelas equações dadas a seguir:

*Equação das observações:*

$$p(y_t / \alpha_1, \dots, \alpha_t, y_1, \dots, y_{t-1}) = p(y_t / Z_t \alpha_t).$$

(2.2)

Os vetores de estado são determinados pela seguinte relação:

*Equação do estado:*

$$\alpha_{t+1} = T_t \alpha_t + R_t \eta_t \quad \eta_t \sim p(\eta_t). \quad (2.3)$$

para  $t = 1, \dots, n$ . Os  $\eta_t$ 's são serialmente independentes e  $p(y_t/Z_t \alpha_t)$  e/ou  $p(\eta_t)$  podem ser não Gaussianos. A matriz  $Z_t$  tem a mesma finalidade e forma do caso do modelo Gaussiano linear em (2.1). Definimos como sinal,  $\theta_t$ , o termo  $Z_t \alpha_t$ . Começamos considerando a forma geral de  $p(y_t/\theta_t)$  particularizando para o caso de distribuições da família exponencial com densidades na forma:

$$p(y_t/\theta_t) = \exp[y_t' \theta_t - b_t(\theta_t) + c_t(y_t)], \quad -\infty < \theta_t < \infty, \quad (2.4)$$

onde  $b_t(\theta_t)$  é duas vezes diferenciável e  $c_t(y_t)$  é apenas função de  $y_t$ . O modelo (2.4) junto com (2.3) é conhecido como *dynamic generalised linear model*[46]. O nome é proveniente do tratamento de (2.4) fora do ambiente de série temporais, com  $\theta_t = Z_t \alpha$  onde  $\alpha$  não depende do tempo, conhecido como *generalised linear model*[32].

Para o modelo (2.4), temos:

$$\begin{aligned} \dot{b}_t(\theta_t) &= \frac{\partial b_t(\theta_t)}{\partial \theta_t}. \\ \ddot{b}_t(\theta_t) &= \frac{\partial^2 b_t(\theta_t)}{\partial \theta_t \partial \theta_t'}. \end{aligned} \quad (2.5)$$

Vamos abreviar  $\dot{b}_t(\theta_t)$  por  $\dot{b}_t$  e  $\ddot{b}_t(\theta_t)$  por  $\ddot{b}_t$  onde não há necessidade de enfatizar a dependência em  $\theta_t$ . Usando os resultados

$$\begin{aligned} E \left[ \frac{\partial \log p(y_t/\theta_t)}{\partial \theta_t} \right] &= 0, \\ E \left[ \frac{\partial^2 \log p(y_t/\theta_t)}{\partial \theta_t \partial \theta_t'} \right] + E \left[ \left( \frac{\partial \log p(y_t/\theta_t)}{\partial \theta_t} \right)^2 \right] &= 0, \end{aligned} \quad (2.6)$$

derivamos de (2.4) e (2.6) que

$$E[y_t] = \dot{b}_t, \quad Var[y_t] = \ddot{b}_t.$$

Deste modo,  $\ddot{b}_t$  deve positivo definido para modelos não-degenerados.

Os resultados de (2.6) são válidos assumindo certas condições de regularidade. Diferenciando a relação

$$\int p(y_t/\theta_t) dy_t = 1,$$

duas vezes com respeito a  $\theta_t$  e usando o resultado

$$\frac{\partial \log p(\cdot)}{\partial \theta_t} = [p(\cdot)]^{-1} \frac{\partial p(\cdot)}{\partial \theta_t}.$$

Particularizando a forma geral de  $p(y_t/\theta_t)$  para o caso onde as observações são geradas pela seguinte relação:

$$y_t = \theta_t + \varepsilon_t, \quad \varepsilon_t \sim p(\varepsilon_t), \tag{2.7}$$

onde os  $\varepsilon_t$ 's são não Gaussianos e serialmente independentes. Podemos exemplificar modelando  $\varepsilon_t$  por uma distribuição *t-Student*. Modelando a partir de (2.7) temos que a equação do log da densidade como:

$$\log p(\varepsilon_t) = \log a(v) + \frac{1}{2} \log \lambda - \frac{v+1}{2} \log(1 + \lambda \varepsilon_t^2), \tag{2.8}$$

onde  $v$  é o número de graus de liberdade e

$$a(v) = \frac{\Gamma(\frac{v}{2} + \frac{1}{2})}{\Gamma(\frac{v}{2})}, \quad \lambda^{-1} = (v-2)\sigma_\varepsilon^2, \quad \sigma_\varepsilon^2 = Var(\varepsilon_t), \quad v > 2, \quad t = 1, \dots, n. \tag{2.9}$$

A média de  $\varepsilon_t$  é zero e a variância é  $\sigma_\varepsilon^2$  para qualquer grau de liberdade  $v$ , que não precisa ser número inteiro. As quantidades  $\varepsilon_t$  e  $\sigma_\varepsilon^2$  podem variar no tempo fazendo com que  $\lambda$  também varie.

### 2.3.2

#### Modelos não lineares

Nesta subsecção apresenta-se uma classe de modelos não lineares obtida do modelo gaussiano linear em (2.1). Permitindo que  $y_t$  dependa não linearmente de  $\alpha_t$  na equação da observação e  $\alpha_{t+1}$  dependa não linearmente de  $\alpha_t$  na equação do estado. Temos então o modelo:

-Equação das observações:

$$y_t = Z_t(\alpha_t) + \epsilon_t, \quad \epsilon_t \sim N(0, H_t). \quad (2.10)$$

-Equação do estado:

$$\alpha_{t+1} = T_t(\alpha_t) + R_t\eta_t, \quad \eta_t \sim N(0, Q_t). \quad (2.11)$$

para  $t = 1, \dots, n$  onde  $Z_t(\cdot)$  e  $T_t(\cdot)$  são vetores diferenciáveis de  $\alpha_t$  com dimensões  $p$  e  $m$  respectivamente. Podemos estender este modelo fazendo  $\epsilon_t$  e  $\eta_t$  serem não Gaussianos[1].

Um exemplo da relação (2.10) é o modelo SV simples. Aqui, o retorno do ativo é uma combinação multiplicativa das componentes gerando:

$$r_t = \bar{\sigma} e^{0.5h_t} \xi_t, \quad \xi_t \sim N(0, 1),$$

$$h_t = \phi h_{t-1} + \eta_t, \quad \eta_t \sim N(0, \sigma_\eta^2), \quad |\phi| < 1.$$

Onde  $r_t$  é o retorno do ativo;  $\bar{\sigma}$  o nível de volatilidade média (fator de escala) e  $\phi$  um coeficiente autoregressivo. Este modelo, que é o modelo de estudo desta dissertação, será apresentado detalhadamente no próximo capítulo.

### 2.3.3

#### Estimação dos hiperparâmetros dos modelos não Gaussianos e/ou não lineares via metodologia Durbin e Koopman

No caso linear Gaussiano, as densidades *a priori* e *a posteriori* possuem forma fechada, ou seja, forma analítica. No caso não linear e/ou não Gaussiano geralmente estas integrais não possuem forma fechada. Desta maneira, necessitam de tratamento numérico ou algum tipo de aproximação.

Utilizaremos a metodologia Durbin e Koopman[13], na qual parte da verossimilhança é obtida através de uma aproximação linear/Gaussiana e a outra via simulação de Monte Carlo, constituindo a MCL. A MCL é baseada nas técnicas de Amostragem por Importância[36] com o uso de variáveis antitéticas[13, 18].

Teoricamente, pode-se sortear uma amostra aleatória da distribuição  $p(\alpha/y, \phi)$  e a partir daí, extrair os momentos da distribuição. Na prática, em se tratando de modelos não Gaussianos/não lineares, a expressão  $p(\alpha/y, \phi)$  não está disponível. Por isso, busca-se uma densidade mais próxima possível de  $p(\alpha/y, \phi)$  com objetivo de obter amostras aleatórias da densidade de interesse. Esta técnica é conhecida como Amostragem por Importância e a densidade mencionada chama-se densidade de importância. A utilização da densidade de importância está explicitada, no final da subsecção, na descrição do procedimento de estimação.

As variáveis antitéticas são incorporadas com a finalidade de melhorar a eficiência computacional no cálculo das integrais sobre dois aspectos: redução do número de amostras no procedimento do suavizador por simulação e redução da variância do estimador.

Durbin e Koopman[13] aproximam o cálculo da verossimilhança para um modelo Gaussiano que pode ser calculado a partir do filtro de Kalman, sendo corrigido via simulação. O cálculo da verossimilhança é exato dado que o erro de estimação é devido ao erro de simulação.

A verossimilhança  $L(y/\psi)$  pode ser definida por  $L(y/\psi) = p(y/\psi)$ , representada pela equação

$$L(y/\psi) = \int p(\alpha, y/\psi) d\alpha. \quad (2.12)$$

Dividindo e multiplicando pela densidade condicional Gaussiana aproximada do vetor de estado, conhecida como densidade de importância<sup>2</sup>

<sup>2</sup>Detalhes em Durbin & Koopman (2001) página 190.

$p_G(\alpha/y, \psi)$ , temos

$$L(y/\psi) = \int \frac{p_{true}(\alpha, y/\psi)}{p_G(\alpha/y, \psi)} p_G(\alpha/y, \psi) d\alpha. \quad (2.13)$$

Seja

$$p_G(\alpha/y, \psi) = \frac{p_G(\alpha, y/\psi)}{p_G(y/\psi)},$$

então

$$\begin{aligned} L(y/\psi) &= p_G(y/\psi) \int \frac{p_{true}(\alpha, y/\psi)}{p_G(\alpha, y/\psi)} p_G(\alpha/y, \psi) d\alpha, \\ L(y/\psi) &= L_G(y/\psi) E_G \left[ \frac{p_{true}(\alpha, y/\psi)}{p_G(\alpha, y/\psi)} \right]. \end{aligned} \quad (2.14)$$

Na expressão acima,  $L_G(y/\psi) = p_G(y/\psi)$  é a verossimilhança do modelo linear Gaussiano aproximado empregado para obter a densidade de importância  $p_G(\alpha/y, \psi)$  e  $E_G$  indica a esperança com respeito a densidade  $p_G(\alpha/y, \psi)$ . Sabendo que o vetor  $\alpha$  depende de  $\alpha_1$  e  $\eta$ , onde  $\alpha_1$  é conhecido, substituímos  $\alpha$  por  $\eta$  na equação (2.14) com propósito de diminuir a dimensionalidade da integração. A equação (2.14) pode ser expressa da forma:

$$L(y/\psi) = L_G(y/\psi) E_G \left[ \frac{p_{true}(\eta, y/\psi)}{p_G(\eta, y/\psi)} \right]. \quad (2.15)$$

A partir de (2.14), dado que

$$\frac{p_{true}(\alpha, y/\psi)}{p_G(\alpha, y/\psi)} = \frac{p_{true}(\alpha/\psi) p_{true}(y/\alpha, \psi)}{p_G(\alpha/\psi) p_G(y/\alpha, \psi)},$$

quando o estado é linear e Gaussiano,  $p_{true}(\alpha/\psi) = p_G(\alpha/\psi)$ , temos que

$$\frac{p_{true}(\alpha, y/\psi)}{p_G(\alpha, y/\psi)} = \frac{p_{true}(y/\alpha, \psi)}{p_G(y/\alpha, \psi)} = \frac{p_{true}(y/\theta, \psi)}{p_G(y/\theta, \psi)}.$$

Onde  $\theta$  é o vetor de sinais  $\theta_t = Z\alpha_t$ . Se a equação das observações tem estrutura do tipo  $y_t = \theta_t + \epsilon_t$ , como em nosso caso, então a verossimilhança é dada por

$$L(y/\psi) = L_G(y/\psi) E_G \left[ \frac{p_{true}(\epsilon/\psi)}{p_G(\epsilon/\psi)} \right]. \quad (2.16)$$

Tirando o logaritmo da verossimilhança, obtemos:

$$\ln L(y/\psi) = \ln L_G(y/\psi) + \ln \bar{w}. \quad (2.17)$$

Sendo

$$\bar{w} = E_G \left[ \frac{p_{true}(\epsilon/\psi)}{p_G(\epsilon/\psi)} \right].$$

$y = (y_1, y_2, \dots, y_T)$  e  $\epsilon = (\epsilon_1, \epsilon_2, \dots, \epsilon_T)$  são os vetores das observações e do distúrbio das medidas respectivamente;  $\ln L_G(y/\psi)$  é o log da verossimilhança do modelo Gaussiano aproximado;  $p_{true}(\epsilon/\psi)$  é a função de densidade dos distúrbios das medidas ;  $p_G(\epsilon/\psi)$  é a densidade Gaussiana do distúrbio das medidas aproximada do modelo e  $\bar{w}$  é a média do fator de escala  $w$ , apresentados mais adiante.

O que torna o modelo elegante do ponto de vista da estimação é o fato de que o log da verossimilhança do modelo não Gaussiano pode ser expresso como o log da verossimilhança do modelo Gaussiano aproximado mais uma correção, estimada por simulação, a qual expressa o distanciamento da suposição Gaussiana em relação ao modelo verdadeiro.

Abaixo apresentamos, a descrição do procedimento de estimação do modelo SV, utilizando Amostragem por Importância e variáveis antitéticas, divididos em sete passos[39]:

1. Escolher um modelo Gaussiano aproximado,  $p_G(\epsilon/y, \psi)$ , de modo que a moda da densidade Gaussiana aproximada( densidade por importância) seja aproximadamente igual a moda da distribuição verdadeira. Para tal, é necessário encontrar média e variância do distúrbio da equação das observações no sentido de aproximar ao máximos estas densidades. Fixando a média em zero, determinar a variância  $H_t$ , onde  $\epsilon_t \sim N(0, H_t)$ , via FKS.
2. Calcular e guardar  $\ln L_G(y/\psi)$  e  $\hat{\epsilon} = E[\epsilon/y, \psi]$  para o modelo aproximado via FKS.
3. Gerar uma amostra  $\epsilon^{(i)} = (\epsilon_1^{(i)}, \dots, \epsilon_T^{(i)})$  da densidade de importância  $p_G(\epsilon/y, \psi)$  do passo 1. Uma versão particular do suavizador por simulação foi apresentada em §2.2.3.
4. Construir amostra antitética:  $\tilde{\epsilon}^{(i)}, \acute{\epsilon}^{(i)}, \grave{\epsilon}^{(i)}, \check{\epsilon}^{(i)}$ .

5. Calcular e armazenar

$$w^{(i)} = \frac{w(\tilde{\epsilon}^{(i)}) + w(\epsilon^{(i)}) + w(\hat{\epsilon}^{(i)}) + w(\check{\epsilon}^{(i)})}{4}, \quad w(\epsilon) = \frac{p_{true}(\epsilon/\psi)}{p_G(\epsilon/\psi)}.$$

6. Repetir os passos 3-5 até sortear  $N$  amostras de  $w^{(i)}$ .

7. Calcular  $\bar{w}$  e  $s_w^2$ , que são média e variância de  $w^{(i)}, i = 1, \dots, N$ :

$$\bar{w} = \frac{1}{N} \sum_{i=1}^N w_i \quad s_w^2 = \frac{1}{N-1} \sum_{i=1}^N (w_i - \bar{w})^2.$$

A partir do cálculo de  $\bar{w}$  e  $s_w^2$  pode-se estimar os hiperparâmetros,  $\psi$ , otimizando numericamente o log da verossimilhança apresentada em (2.18). Os valores iniciais podem vir da log verossimilhança aproximada  $\ln L_G(y/\psi)$ . A escolha do número de simulações determinará a acurácia da estimação.